**Characterization of the non-coding control region of polyomavirus KI isolated from nasopharyngeal samples from patients with respiratory symptoms or infection and blood from healthy blood donors in Norway**

Xiaobo Song1,§, Marijke Van Ghelue2,3,§, Maria Ludvigsen1, Svein Arne Nordbø4,5, Bernhard Ehlers6, Ugo Moens1,\*

1University of Tromsø, Faculty of Health Sciences, Institute of Medical Biology, NO-9037 Tromsø, Norway; 2University Hospital of North Norway, Department of Medical Genetics, NO-9038 Tromsø, Norway; 3University of Tromsø, Faculty of Health Sciences, Institute of Clinical Biology, NO-9037 Tromsø, Norway; 4Department of Medical Microbiology, Trondheim University Hospital, NO-7489 Trondheim, Norway; 5Institute of Laboratory Medicine, Children’s and Women’s Health, Norwegian University of Science and Technology, Trondheim, Norway.6”Division 12 Measles, Mumps, Rubella and Viruses Affecting Immunocompromised Patients“, Robert Koch Institute, Berlin, Germany.

§These authors contributed equally to this work.

\*Corresponding authors: E-mail: ugo.moens@uit.no; Tel: +47-77644622; Fax: +47-77645350

Running title: promoter activity KI polyomavirus variants

**Abstract**

Seroepidemiological studies showed that the human polyomavirus KI (KIPyV) is common in the human population with age-specific seroprevalence ranging from 40-90%. Genome epidemiological analyses demonstrated that KIPyV DNA is predominantly found in respiratory tract samples of immunocompromised individuals and children suffering from respiratory diseases, but viral sequences have also been detected in brain, tonsil, lymphoid tissue studies, plasma, blood, and faeces. Little is known about the sequence variation in the non-coding control region of KIPyV variants residing in different sites of the human body and whether specific strains dominate in certain parts of the world. In this study, we sequenced the non-coding control region (NCCR) of naturally occurring KIPyV variants in nasopharyngeal samples from paediatric patients with respiratory symptoms or infection and in blood from healthy donors in Norway. In total 86 sequences were obtained, 44 of which were identical to the original isolated Stockholm 60 variant. The remaining NCCRs contained one or several mutations, none of them previously reported. The same mutations could be detected in NCCRs amplified from blood and nasopharyngeal samples. Some patients had different variants in their specimens. Transient transfection studies in HEK293 cell with a luciferase reporter plasmid demonstrated that some single mutations had a significant effect on the relative early and late promoter strength compared to Stockholm 60 promoter. The effect of the NCCR mutations on viral replication and possible virulence properties remains to be established.

Keywords: polyomavirus, mutation, blood, nasopharyngeal samples, transcriptional control region, luciferase assay

**Introduction**

Polyomaviruses are non-enveloped viruses with a circular dsDNA genome of approximately 5,000 base-pairs. These viruses are common in birds and mammals and recently the first complete fish polyomavirus genome has been published ((Johne et al., 2011);(Peretti et al., 2015)). Their genome can be divided into three functional regions: the early region encoding regulatory proteins required for transcription and viral DNA replication, the late region encoding the capsid proteins, and interspersed between these the non-coding control region (NCCR) encompassing the origin of replication and the promoters controlling transcription of the early and late genes, respectively ((DeCaprio and Garcea, 2013)). The first human polyomaviruses, BK and JC, were detected using histology or cytology followed by electron microscopy, and their existence was reported in 1971 ((Gardner et al., 1971); (Padgett et al., 1971)). The development of new techniques such as high throughput sequencing, rolling circle amplification and polymerase chain reaction with degenerated primers has led in recent years to the discovery of new polyomaviruses that can infect humans. KI polyomavirus, named after the researchers’ institute (Karolinska Institute) was the first novel human polyomavirus to be characterized in 2007 ((Allander et al., 2007)). Since then, 10 additional human polyomaviruses have been identified: WUPyV, Merkel cell polyomavirus (MCPyV), HPyV6, 7, 9, 10, 12, Trichodysplasia spinulosa-associated polyomavirus (TSPyV), STL-PyV (and the isolates MW and MX), and NJPyV-2013 ((Gaynor et al., 2007); (Feng et al., 2008); (Schowalter et al., 2010); (van der Meijden et al., 2010); (Scuda et al., 2011); (Buck et al., 2012); (Siebrasse et al., 2012); (Yu et al., 2012); (Korup et al., 2013); (Lim et al., 2013); (Mishra et al., 2014)).

KIPyV was originally isolated from 20 pooled nasopharyngeal aspirates (Allander et al., 2007). Seroepidemiological studies revealed that KIPyV infection is common in the human population with seroprevalence >50% in the adult age group. ((Kean et al., 2009); (Nguyen et al., 2009); (Kantola et al., 2010); (Neske et al., 2010); (Gossai et al., 2016)). Primary infection seems to occur in early childhood as antibodies against the capsid protein VP1 can already be detected in babies <6 month old (Nguyen et al., 2009), but there is no evidence for mother-to-foetus transmission of the virus (Sadeghi et al., 2010). Little is known about the cell tropism of this virus, but viral sequences are predominantly found in respiratory tract secretions with prevalence ranging between 0.5-6.5% (Babakir-Mina et al., 2013). KIPyV DNA has occasionally been amplified from paranasal tissue (Babakir-Mina et al., 2009b), tonsil (Babakir-Mina et al., 2009b),(Astegiano et al., 2010), lymphoid tissue (Sharp et al., 2009), lung tissue (Babakir-Mina et al., 2009c),(Teramoto et al., 2011), stool (Allander et al., 2007),(Babakir-Mina et al., 2009a),(Bialasiewicz et al., 2009),(Kantola et al., 2009),(Mourez et al., 2009),(Li et al., 2013), brain (Barzon et al., 2009b), eyebrow hair (Hampras et al., 2015), normal skin (Hampras et al., 2015), and blood and plasma (Barzon et al., 2009a),(Babakir-Mina et al., 2010),(Csoma et al., 2012),(Touinssi et al., 2013). Immunohistochemical assay with a monoclonal antibody against the capsid protein VP1 of KIPyV of spleen tissue from a 42-year-old HIV-positive male stained positive, but the identity of the KIPyV VP1 positive cell type could not be determined (Siebrasse et al., 2014).

Little information is available on genome sequence variability of circulating KIPyV strains in different biological specimens and how mutations may affect viral activity remains unknown. So far, the complete genome sequence of 10 KIPyV strains have been deposited in the GenBank (Stockholm variants 60, 350, 380; Brisbane variants 001, 002, 003; CU-255 and -258; FZ52, and HN057), while the NCCR sequence of three additional isolates (Brisbane variants 004, 006 and 007) are also available. The KIPyV isolates are derived from Sweden, Australia, China, and Thailand and their NCCRs contain minor changes (**Table 1** and **Supplementary Table S1**). The effect of mutations on the transcriptional activity of the KIPyV early and late promoters has not been examined. In this study, we investigated whether KIPyV strains with different NCCRs are circulating in the Norwegian population and whether polymorphisms in the NCCR affect the early and late promoter activity.

**Results**

*KIPyV detection in NPA and blood*

Nasopharyngeal samples (n=94) from individuals with respiratory symptoms or infection and blood samples (n=96) from healthy blood donors were examined for the presence of KIPyV DNA using previously published primers that amplify part of the VP1 gene (Allander et al., 2007). Each sample was tested three times and a sample was considered positive when minimum 2 out of three PCR gave a positive product. Amplified KIPyV VP1 DNA was obtained for 10 (10.6%) of the nasopharyngeal samples and for 12 (12.5%) of the blood samples.

Next, we wanted to determine whether mutations occurred in the NCCR and whether unique NCCRs were associated with clinical conditions. For this, the complete NCCR was amplified from KIPyV DNA-positive samples. To analyse if several KIPyV genome variants were present in the same sample extract, PCR products were cloned and 3-6 colonies per cloning were sequenced. In total, 86 readable sequences were obtained from 22 subjects (12 patients suffering of respiratory tract disease or infection and 10 blood donors). Forty-four (51%) of the NCCR sequences were completely identical to the NCCR of the previously described Stockholm 60 (EF127906) strain (Allander et al., 2007). The sequence of the Stockholm 60 NCCR is given in **Supplementary Figure 1**. The NCCR of this variant is identical to the NCCR of Brisbane 002, Brisbane 004, Brisbane 006, Brisbane 007, and CU-258 [Supplementary table S1; (Bialasiewicz et al., 2007),(Bialasiewicz et al., 2009), (Payungporn et al., 2008)]. The exuberance of the Stockholm 60 NCCR in biological samples indicates that this NCCR may represent the archetypal NCCR. The NCCR of the remaining 42 naturally occurring KIPyV isolates from blood (hereafter referred to as Bld) and nasopharyngeal samples (hereafter referred to as NPA) displayed minor changes. The mutations are summarized in **Table 1**, and are numbered with nucleotide 1 as the first nucleotide in the NCCR in the early-late orientation. The sequences of the mutant NCCRs were deposited in GenBank (accession numbers KU564911-564952). Twenty-four variants had a single point mutation. Double or triple point mutations were observed in 12 and 5 samples, respectively, while one variant (Bld61c) displayed five point mutations (Supplementary Table S3). One NCCR had a single nucleotide deletion (G352 in NPA57b), while no insertions were found in any of the NCCR variants. The sequences obtained from different clones of the same specimen revealed different mutations, suggesting that different strains may circulate in the same patient. Identical mutations were found in samples from different patients. The mutation T154C was present in samples NPA74c and NPA81c, substitution T28C occurred in NCCR amplified from NPA7a, Bld29b, and Bld60c, A173G was present in the NCCR from NPA81b and Bld61c. The substitution G484A occurred only in NCCR amplified from nasopharyngeal samples (NPA1a, NPA19a, NPA69b, NPA74d, NPA74f, and NPA81a). The NCCR of samples NPA1a and NPA74d (respectively NPA74c and NPA81c) were identical, illustrating that the same variant can circulate in different individuals. To ensure that the mutations did not originate from DNA polymerase-induced mistakes, PCR was performed on the luciferase reporter plasmid containing the Stockholm 60 early promoter (Moens et al., 2015) using the same PCR conditions as applied for the NPA and blood samples. The amplicon was cloned and six colonies were sequenced. All sequences were identical to the Stockholm 60 sequence, suggesting that the PCR and cloning had not introduced unwanted mutations.

*Effect of mutations on early and late promoter activity*

The host tropism of KIPyV is not known and cell culture models sustaining viral replication are lacking. We have previously investigated the early and late promoter activities of all 13 known human polyomaviruses in 10 different cell lines and found that the KIPyV early and late promoters had highest activity in the human embryonal kidney cell line HEK293 (Moens et al., 2015). These cells were therefore used to compare the relative promoter strength of the mutant NCCRs with the archetypal consensus sequence. We decided to test the effect of single mutations on the relative promoter activity because this may allow us to correlate this particular mutation to changed promoter strength. Analysing the effect of multiple mutations on promoter strength does not allow us to pinpoint which mutation(s) contribute(s) to altered activity. Eighteen of 24 isolates containing a single nucleotide substitution were selected and their early and late promoter activities were compared to those of Stockholm 60. These NCCR variants were primarily selected because the ALGGEN PROMO algorithm (Messeguer et al., 2002),(Farre et al., 2003) predicts that the mutations affect putative transcription factor binding sites (**Supplementary** **Table S3**). Investigating single mutations allowed us to speculate on the biological importance of the putative binding sites for transcription factors on the viral promoter activity. Mutations in early and late Bld29d, Bld38d, Bld69c, NPA1a, and NPA7b promoters, and in early Bld62c and NPA74c promoters did not significantly altered the strength compared to the corresponding promoter of Stockholm 60 (**Figure 1**). Mutations in NCCR variants Bld22c, Bld45d, and Bld71c resulted in reproducible tendency towards reduced early and late activity. A significant (p<0.0002) reduction was observed for the NPA89d and NPA44a early and late promoters, the Bld38c and Bld60c early promoters, and the NPA57a and NPA74c late promoters. A significant increase was measured for the early and late NPA7d and Bld60d (p<0.001) promoters. Most mutations resulted in a tendency to increase (respectively decrease) both the early and late promoter activity. Exceptions were the NCCR variants Bld38c (C85T) and NPA94a (T189C), where the early promoter activity decreased and the late promoter activity increased compared to the Stockholm 60 NCCR. The opposite was observed for the Bld29d (A33G) promoter; the early promoter was up to 30% stronger than the early Stockholm 60 promoter (p=0.04), while the late promoter was up to 25% weaker than the late Stockholm 60 promoter (p=0.0007).

*Comparison of the relative Stockholm 60 and Brisbane 001 promoter strength in HEK293 cells*

The NCCRs of Brisbane 001, Brisbane 005, and CU-255 contain the 10 base-pair insertion AGGCGCTGCG between nucleotides A132 and G133 and three additional point mutations (T180C, T183A, C306A) compared to Stockholm 60 (**Table 1 and Supplementary Table S1**; (Bialasiewicz et al., 2007),(Bialasiewicz et al., 2009),(Payungporn et al., 2008)). All three variants were isolated from NPA (Bialasiewicz et al., 2007),(Bialasiewicz et al., 2009),(Payungporn et al., 2008). Because of the 10 base-pair insertion unique to these isolates, we decided to compare the relative strength of the Brisbane 001 early and late promoter to Stockholm 60 in HEK293 cells. Both the early and late promoter of Brisbane 001 were significantly (p<0.001) stronger than the corresponding promoters of Stockholm 60 (**Figure 2**). The early Brisbane 001 promoter was on average 1.8-fold stronger, while the late promoter was 1.3-fold stronger than the corresponding Stockholm 60 promoter.

**Discussion**

PCR-based studies on respiratory tract samples collected from different patient groups in all ages and different parts of the world showed a KIPyV DNA prevalence ranging from 0.5-12.1% (Babakir-Mina et al., 2013; Iaria et al., 2015). On the other hand, longitudinal studies on nasal swabs from 56 healthy children in Australia during their first 18 months of life and on haematopoietic stem cell recipients showed that respectively 45% and 26% of subjects had >1 KIPyV DNA positive sample (Kuypers et al., 2012; Rockett et al., 2015). KIPyV DNA has been detected in blood from healthy individuals, HIV-positive patients and renal transplant patients with a frequency of 0-3.2% (Allander et al., 2007; (Bialasiewicz et al., 2008); Barzon, Squarzon, Militello, Trevisan, & Palu, 2009; Csoma, Meszaros, Asztalos, Konya, & Gergely, 2011)(Babakir-Mina et al., 2013). We found that 10.6% of the nasopharyngeal samples (n=94) taken from patients with respiratory symptoms or infection contained KIPyV DNA and 12.5% (n=96) of the blood specimens from healthy individuals were positive for VP1 DNA. These prevalences are higher than found by other groups, but this could be due to the patient group examined. Upper respiratory infections may predispose the host to higher susceptibility for KIPyV infection or reactivation. The use of different viral DNA extraction methods, primers and PCR conditions, geographic differences, and age of the subjects may also affect the genoprevalence.

About 50% of the NCCR sequences from our samples were identical with the NCCR of Stockholm 60, suggesting that this strain may represent the archetypal strain. This is underscored by the finding of KIPyV with Stockholm 60 NCCR in individuals from Australia, Thailand and China (**Supplementary Table S1**). Fifty-seven percent of the NCCRs possessed single nucleotide substitutions, while 43% had two or more point mutations. None of the mutations identified in this cohort were reported in other studies, while previously reported mutations were not found here, in the NCCR sequences from NPA and blood of Norwegian patients. This may suggest that specific variants may circulate in different geographic regions. However, additional sequences must be obtained to demonstrate a possible geographic distribution of KIPyV subtypes. One mutation (G484A) was only detected in NCCRs amplified from nasopharyngeal samples, suggesting that this mutation may be specific for KIPyV circulating in the respiratory system. Viruses with different NCCR were detected in the same sample of a particular patient, underscoring that viral replication may have occurred. Indeed PyV replication leads to mutations in the NCCR (Imperiale and Jiang, 2015). However, the genuine host cell(s) for KIPyV remain(s) to be identified, and transcription factors involved in early and late transcription have not been characterized. To explore the impact of NCCR polymorphisms on promoter activity, we initiated studies with some of the NCCRs carrying a single mutation. Most mutations altered the strength of the viral promoter activity, but it can only be speculated what the effect on binding of transcription factors is of the mutations reported in this study. Using the ALGGEN-PROMO algorithm (Messeguer et al., 2002),(Farre et al., 2003) we examined the effect of mutations on the putative transcription factor binding sites. Most point mutations removed or introduced multiple possible binding sites. However, the mutations A57G (NPA7b), A166G (NPA7d), C208T (Bld62c), G291A (Bld71c), and G484A (NPA1a) remove a single putative site for respectively Pu box binding factor, c-Myb, E2F-1, AP2A, and GATA-2 site. The T154C substitution (NPA74c) generates a putative Pax 5 binding site (**Supplementary Table S3**). Co-transfection studies with a c-Myb expression plasmid showed that both the early and late NPA7a promoter were still activated with levels similar to the Stockholm 60 early and late promoter (results not shown). These observations indicate that despite the usefulness of transcription factor site algorithms, experimental studies are absolutely required to confirm the functionality of a putative binding site.

In conclusion, our results suggest that Stockholm 60 is the archetypal KIPyV strain and that small changes (point mutations) are present in strains circulating in the human population. However, larger rearrangements as seen for the BKPyV and JCPyV promoters seem to be rare. The absence of a cell culture system that sustains replication of KIPyV and the lacking association of KIPyV with disease does not allow conclusions on the effect of these NCCR variants on viral replication and virulence.

**Materials and Methods**

*Clinical samples*

Blood samples (n=96) were collected from healthy blood donors, while 94 nasopharyngeal samples were obtained from patients with respiratory symptoms or infection. This study was approved by the Ethics Committee of the University of Tromsø (REK2012/410). Viral DNA was extracted from nasopharyngeal samples by QIAamp MinElute Virus Spin Kit (Qiagen, cat. no. 57704), while DNA from whole blood was extracted isolated by the use of an automated DNA extractor Genovision M48 (QIAGEN, Hilden, Germany).

*PCR and cloning*

Nasopharyngeal and blood samples were originally screened for the presence of KIPyV DNA using the primers 5’-CGCAGTACCACTGTCAGAAGAAC-3’ and 5’-TTCTGCCAGGCTGTAACATAC-3’, amplifying a fragment of the VP1 gene (Allander et al., 2007). KIPyV NCCR was amplified from VP1-positive samples using the primer set: 5'-CGCAGTACCACTGTCAGAAGAAAC-3' and 5'-AGATCTCTAAAAACAAAATAAAAATTGGTGACCCT-3'. PCR conditions were 5 min at 95oC, followed by 40 cycles of 30 sec at 95oC, 1 min at 54oC, and 1 min at 72oC. A final extension step of 10 min at 72oC was included. The AccuPrimeTM Taq DNA polymerase, high fidelity with 3’→5’ exonuclease activity (proofreading) was used (ThermoFisher Scientific, cat. No. 12346-086). The PCR products were cloned and transformed in competent *E. coli* DH5 cells. After transformation, several colonies were picked, plasmid DNA was purified (Machery-Nagel NucleoSpin Plasmid cat. no. 740588.250) and sequenced. Cycle sequencing reactions were performed using Big Dye 3.1 Sequencing kit (Life Technologies, cat. no. 4337455).

*Plasmids*

The plasmids containing the KIPyV strain Stockholm 60 early and late promoter in the pGL3-basic luciferase reporter plasmid have been previously described (Moens et al., 2015). These plasmids were used to introduce single nucleotide substitutions using the primers given in **Supplementary Table S2**. The reporter plasmids with the Brisbane 001 early and late promoter respectively were generated by site-directed mutagenesis of the aforementioned plasmids by first using a primer set that resulted in addition of 10 bp. Subsequently, two additional site-directed mutagenesis reactions were performed to introduce the point mutations (See Supplementary Table S2 for the sequences of the site-directed mutagenesis primers). All mutations were verified by sequencing.

*Cells*

HEK293 cells were maintained in Dulbecco’s Modified Eagle’s Medium (Sigma D5796) with 10% foetal bovine serum (Life Technologies, cat. no. 10500-064).

*Transfection*

Cells were seeded out in twelve-well cell culture plates and transfected the following day. Cells were approximately 70-85% confluent at the day of transfection. JetPrime® (Polyplus, Illkirch, France) was used as transfection reagent according to the manufacturer’s protocol. Each transfection was repeated 3-5 times with three independent parallels each.

*Luciferase assays*

Cells were lysed approximately 24 h post-transfection in 100 l Luciferase Assay Tropix Lysis solution with 0.5 mM DTT freshly added. Cells were scraped and transferred to Eppendorf tubes, followed by 3 min centrifugation at 12,000g. Supernatant (20 l) was transferred to 96-well microtiter plate and luciferase buffer (Promega) was added. Light units were measured in a luminometer (Labsystem Luminoscan RT). The activity of the early (respectively late) Stockholm 60 promoter was arbitrary set as 100 and the activity of the promoters of the other variants was related to this. Luciferase values were adjusted for the protein concentration in each sample. The protein concentration was measured using the Protein Quantification Assay from Macherey-Nagel (Düren, Germany).

*Statistical analysis*

The *t*-test was employed to determine statistical differences between the Stockholm 60 and mutant promoters.

**References**

**Allander, T., Andreasson, K., Gupta, S., Bjerkner, A., Bogdanovic, G., Persson, M.A., Dalianis, T., Ramqvist, T., & Andersson, B., (2007).** Identification of a third human polyomavirus. *J Virol* **81,** 4130-4136.

**Astegiano, S., Terlizzi, M.E., Elia, M., Cavallo, G.P., Costa, C., Cavallo, R., & Bergallo, M., (2010).** Prevalence of polyomaviruses BK, JC, SV40, KI, and WU in non-malignant tonsil specimens. *Minerva Med* **101,** 385-389.

**Babakir-Mina, M., Ciccozzi, M., Alteri, C., Polchi, P., Picardi, A., Greco, F., Lucarelli, G., Arcese, W., Perno, C.F., & other authors, (2009a).** Excretion of the novel polyomaviruses KI and WU in the stool of patients with hematological disorders. *J Med Virol* **81,** 1668-1673.

**Babakir-Mina, M., Ciccozzi, M., Bonifacio, D., Bergallo, M., Costa, C., Cavallo, R., Di Bonito, L., Perno, C.F., & Ciotti, M., (2009b).** Identification of the novel KI and WU polyomaviruses in human tonsils. *J Clin Virol* **46,** 75-79.

**Babakir-Mina, M., Ciccozzi, M., Campitelli, L., Aquaro, S., Lo Coco, A., Perno, C.F., & Ciotti, M., (2009c).** Identification of the novel KI Polyomavirus in paranasal and lung tissues. *J Med Virol* **81,** 558-561.

**Babakir-Mina, M., Ciccozzi, M., Farchi, F., Bergallo, M., Cavallo, R., Adorno, G., Perno, C.F., & Ciotti, M., (2010).** KI and WU polyomaviruses and CD4+ cell counts in HIV-1-infected patients, Italy. *Emerg Infect Dis* **16,** 1482-1485.

**Babakir-Mina, M., Ciccozzi, M., Perno, C.F., & Ciotti, M., (2013).** The human polyomaviruses KI and WU: virological background and clinical implications. *APMIS* **121,** 746-754.

**Barzon, L., Squarzon, L., Militello, V., Trevisan, M., & Palu, G., (2009a).** Human KI and WU polyomavirus infection in immunocompromised subjects. *J Clin Virol* **45,** 370.

**Barzon, L., Squarzon, L., Militello, V., Trevisan, M., Porzionato, A., Macchi, V., De Caro, R., & Palu, G., (2009b).** WU and KI polyomaviruses in the brains of HIV-positive patients with and without progressive multifocal leukoencephalopathy. *J Infect Dis* **200,** 1755-1758.

**Bialasiewicz, S., Whiley, D.M., Lambert, S.B., Jacob, K., Bletchly, C., Wang, D., Nissen, M.D., & Sloots, T.P., (2008).** Presence of the newly discovered human polyomaviruses KI and WU in Australian patients with acute respiratory tract infection. *J Clin Virol* **41,** 63-68.

**Bialasiewicz, S., Whiley, D.M., Lambert, S.B., Nissen, M.D., & Sloots, T.P., (2009).** Detection of BK, JC, WU, or KI polyomaviruses in faecal, urine, blood, cerebrospinal fluid and respiratory samples. *J Clin Virol* **45,** 249-254.

**Bialasiewicz, S., Whiley, D.M., Lambert, S.B., Wang, D., Nissen, M.D., & Sloots, T.P., (2007).** A newly reported human polyomavirus, KI virus, is present in the respiratory tract of Australian children. *J Clin Virol* **40,** 15-18.

**Buck, C.B., Phan, G.Q., Raiji, M.T., Murphy, P.M., McDermott, D.H., & McBride, A.A., (2012).** Complete genome sequence of a tenth human polyomavirus. *J Virol* **86,** 10887.

**Csoma, E., Sapy, T., Meszaros, B., & Gergely, L., (2012).** Novel human polyomaviruses in pregnancy: higher prevalence of BKPyV, but no WUPyV, KIPyV and HPyV9. *J Clin Virol* **55,** 262-265.

**DeCaprio, J.A., & Garcea, R.L., (2013).** A cornucopia of human polyomaviruses. *Nat Rev Microbiol* **11,** 264-276.

**Farre, D., Roset, R., Huerta, M., Adsuara, J.E., Rosello, L., Alba, M.M., & Messeguer, X., (2003).** Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res* **31,** 3651-3653.

**Feng, H., Shuda, M., Chang, Y., & Moore, P.S., (2008).** Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* **319,** 1096-1100.

**Gardner, S.D., Field, A.M., Coleman, D.V., & Hulme, B., (1971).** New human papovavirus (B.K.) isolated from urine after renal transplantation. *Lancet* **1,** 1253-1257.

**Gaynor, A.M., Nissen, M.D., Whiley, D.M., Mackay, I.M., Lambert, S.B., Wu, G., Brennan, D.C., Storch, G.A., Sloots, T.P., & other authors, (2007).** Identification of a novel polyomavirus from patients with acute respiratory tract infections. *PLoS Pathog* **3,** e64.

**Gossai, A., Waterboer, T., Nelson, H.H., Michel, A., Willhauck-Fleckenstein, M., Farzan, S.F., Hoen, A.G., Christensen, B.C., Kelsey, K.T., & other authors, (2016).** Seroepidemiology of Human Polyomaviruses in a US Population. *Am J Epidemiol* **183,** 61-69.

**Hampras, S.S., Giuliano, A.R., Lin, H.Y., Fisher, K.J., Abrahamsen, M.E., McKay-Chopin, S., Gheit, T., Tommasino, M., & Rollison, D.E., (2015).** Natural History of Polyomaviruses in Men: The HPV Infection in Men (HIM) Study. *J Infect Dis* **211,** 1437-1446.

**Iaria, M., Caccuri, F., Apostoli, P., Giagulli, C., Pelucchi, F., Padoan, R.F., Caruso, A., & Fiorentini, S., (2015).** Detection of KI WU and Merkel cell polyomavirus in respiratory tract of cystic fibrosis patients. *Clin Microbiol Infect* **21,** 603.e609-603.e615.

**Imperiale, M.J., & Jiang, M., (2015).** What DNA viral genomic rearrangements tell us about persistence. *J Virol* **89,** 1948-1950.

**Johne, R., Buck, C.B., Allander, T., Atwood, W.J., Garcea, R.L., Imperiale, M.J., Major, E.O., Ramqvist, T., & Norkin, L.C., (2011).** Taxonomical developments in the family Polyomaviridae. *Arch Virol* **156,** 1627-1634.

**Kantola, K., Sadeghi, M., Ewald, M.J., Weissbrich, B., Allander, T., Lindau, C., Andreasson, K., Lahtinen, A., Kumar, A., & other authors, (2010).** Expression and serological characterization of polyomavirus WUPyV and KIPyV structural proteins. *Viral Immunol* **23,** 385-393.

**Kantola, K., Sadeghi, M., Lahtinen, A., Koskenvuo, M., Aaltonen, L.M., Mottonen, M., Rahiala, J., Saarinen-Pihkala, U., Riikonen, P., & other authors, (2009).** Merkel cell polyomavirus DNA in tumor-free tonsillar tissues and upper respiratory tract samples: implications for respiratory transmission and latency. *J Clin Virol* **45,** 292-295.

**Kean, J.M., Rao, S., Wang, M., & Garcea, R.L., (2009).** Seroepidemiology of human polyomaviruses. *PLoS Pathog* **5,** e1000363.

**Korup, S., Rietscher, J., Calvignac-Spencer, S., Trusch, F., Hofmann, J., Moens, U., Sauer, I., Voigt, S., Schmuck, R., & Ehlers, B., (2013).** Identification of a novel human polyomavirus in organs of the gastrointestinal tract. *PLoS One* **8,** e58021.

**Kuypers, J., Campbell, A.P., Guthrie, K.A., Wright, N.L., Englund, J.A., Corey, L., & Boeckh, M., (2012).** WU and KI polyomaviruses in respiratory samples from allogeneic hematopoietic cell transplant recipients. *Emerg Infect Dis* **18,** 1580-1588.

**Li, K., Guo, J., Zhao, R., Xue, Y., Chen, L., Yang, J., Peng, J., & Jin, Q., (2013).** Prevalence of 10 human polyomaviruses in fecal samples from children with acute gastroenteritis: a case-control study. *J Clin Microbiol* **51,** 3107-3109.

**Lim, E.S., Reyes, A., Antonio, M., Saha, D., Ikumapayi, U.N., Adeyemi, M., Stine, O.C., Skelton, R., Brennan, D.C., & other authors, (2013).** Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. *Virology* **436,** 295-303.

**Messeguer, X., Escudero, R., Farre, D., Nunez, O., Martinez, J., & Alba, M.M., (2002).** PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* **18,** 333-334.

**Mishra, N., Pereira, M., Rhodes, R.H., An, P., Pipas, J.M., Jain, K., Kapoor, A., Briese, T., Faust, P.L., & Lipkin, W.I., (2014).** Identification of a novel polyomavirus in a pancreatic transplant recipient with retinal blindness and vasculitic myopathy. *J Infect Dis* **210,** 1595-1599.

**Moens, U., Van Ghelue, M., Ludvigsen, M., Korup-Schulz, S., & Ehlers, B., (2015).** The early and late promoters of BKPyV, MCPyV, TSPyV, and HPyV12 are among the strongest of all known human polyomaviruses in 10 different cell lines. *J Gen Virol* **96,** 2293-2303.

**Mourez, T., Bergeron, A., Ribaud, P., Scieux, C., de Latour, R.P., Tazi, A., Socie, G., Simon, F., & LeGoff, J., (2009).** Polyomaviruses KI and WU in immunocompromised patients with respiratory disease. *Emerg Infect Dis* **15,** 107-109.

**Neske, F., Prifert, C., Scheiner, B., Ewald, M., Schubert, J., Opitz, A., & Weissbrich, B., (2010).** High prevalence of antibodies against polyomavirus WU, polyomavirus KI, and human bocavirus in German blood donors. *BMC Infect Dis* **10,** 215.

**Nguyen, N.L., Le, B.M., & Wang, D., (2009).** Serologic evidence of frequent human infection with WU and KI polyomaviruses. *Emerg Infect Dis* **15,** 1199-1205.

**Padgett, B.L., Walker, D.L., ZuRhein, G.M., Eckroade, R.J., & Dessel, B.H., (1971).** Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. *Lancet* **1,** 1257-1260.

**Payungporn, S., Chieochansin, T., Thongmee, C., Samransamruajkit, R., Theamboolers, A., & Poovorawan, Y., (2008).** Prevalence and molecular characterization of WU/KI polyomaviruses isolated from pediatric patients with respiratory disease in Thailand. *Virus Res* **135,** 230-236.

**Peretti, A., FitzGerald, P.C., Bliskovsky, V., Pastrana, D.V., & Buck, C.B., (2015).** Genome Sequence of a Fish-Associated Polyomavirus, Black Sea Bass (Centropristis striata) Polyomavirus 1. *Genome Announc* **3,** pii:e01476-14.

**Rockett, R.J., Bialasiewicz, S., Mhango, L., Gaydon, J., Holding, R., Whiley, D.M., Lambert, S.B., Ware, R.S., Nissen, M.D., & other authors, (2015).** Acquisition of human polyomaviruses in the first 18 months of life. *Emerg Infect Dis* **21,** 365-367.

**Sadeghi, M., Riipinen, A., Vaisanen, E., Chen, T., Kantola, K., Surcel, H.M., Karikoski, R., Taskinen, H., Soderlund-Venermo, M., & other authors, (2010).** Newly discovered KI, WU, and Merkel cell polyomaviruses: no evidence of mother-to-fetus transmission. *Virol J* **7,** 251.

**Schowalter, R.M., Pastrana, D.V., Pumphrey, K.A., Moyer, A.L., & Buck, C.B., (2010).** Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. *Cell Host Microbe* **7,** 509-515.

**Scuda, N., Hofmann, J., Calvignac-Spencer, S., Ruprecht, K., Liman, P., Kuhn, J., Hengel, H., & Ehlers, B., (2011).** A novel human polyomavirus closely related to the african green monkey-derived lymphotropic polyomavirus. *J Virol* **85,** 4586-4590.

**Sharp, C.P., Norja, P., Anthony, I., Bell, J.E., & Simmonds, P., (2009).** Reactivation and mutation of newly discovered WU, KI, and Merkel cell carcinoma polyomaviruses in immunosuppressed individuals. *J Infect Dis* **199,** 398-404.

**Siebrasse, E.A., Nguyen, N.L., Smith, C., Simmonds, P., & Wang, D., (2014).** Immunohistochemical detection of KI polyomavirus in lung and spleen. *Virology* **468-470,** 178-184.

**Siebrasse, E.A., Reyes, A., Lim, E.S., Zhao, G., Mkakosya, R.S., Manary, M.J., Gordon, J.I., & Wang, D., (2012).** Identification of MW polyomavirus, a novel polyomavirus in human stool. *J Virol* **86,** 10321-10326.

**Teramoto, S., Kaiho, M., Takano, Y., Endo, R., Kikuta, H., Sawa, H., Ariga, T., & Ishiguro, N., (2011).** Detection of KI polyomavirus and WU polyomavirus DNA by real-time polymerase chain reaction in nasopharyngeal swabs and in normal lung and lung adenocarcinoma tissues. *Microbiol Immunol* **55,** 525-530.

**Touinssi, M., Galicher, V., de Micco, P., & Biagini, P., (2013).** Molecular epidemiology of KI and WU polyomaviruses in healthy blood donors, south-eastern France. *J Med Virol* **85,** 1444-1446.

**van der Meijden, E., Janssens, R.W., Lauber, C., Bouwes Bavinck, J.N., Gorbalenya, A.E., & Feltkamp, M.C., (2010).** Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. *PLoS Pathog* **6,** e1001024.

**Yu, G., Greninger, A.L., Isa, P., Phan, T.G., Martinez, M.A., de la Luz Sanchez, M., Contreras, J.F., Santos-Preciado, J.I., Parsonnet, J., & other authors, (2012).** Discovery of a novel polyomavirus in acute diarrheal samples from children. *PLoS One* **7,** e49449.

**Figure 1.** Effect of single point mutations on the relative early and late promoter activity of KIPyV. HEK293 cells were transfected with a luciferase reporter plasmid containing the early, respectively the late, promoter of KIPyV variants isolated from blood (Bld) or nasopharyngeal aspirates (NPA). Luciferase values were corrected for the protein content in each sample. The early (respectively late) promoter activity of the Stockholm 60 strain was arbitrarily set as 100% and the promoter activities of the variants were related to this. Each box plot represents the average of 3-5 experiments performed with three independent parallels. The box plots of the Stockholm 60 early and late promoter are shown on the right side of the figure. The red line represent the 100% activity of KIPyV Stockholm 60 early (respectively late) promoter. The blue box plots display the early promoter activity, the green plots late promoter activity. The asterisk is a far outlier that showed >1.5 times higher (respectively lower) value than the upper (respectively lower) quartile, while the open dots are outliers that displayed <1.5 times higher (respectively lower) value than the upper (respectively lower) quartile.

**Figure 2.** Comparison of the relative early and late promoter strength of KIPyV strains Stockholm 60 and Brisbane 001. HEK293 cells were transfected with a luciferase reporter containing the early or late promoter of Stockholm 60 (respectively Brisbane 001). The relative promoter activities were calculated as described in the legend of Figure 1. The results represent the average of 3-4 experiments performed with three independent parallels. There was a significant difference (p<0.001) between the early (respectively late) promoter activity of Stockholm 60 and Brisbane 001.

Figure 1



Figure 2



**Table 1.** Mutations in the NCCR of KIPyV isolates compared to the Stockholm60 strain (GenBank accession number EF127906).

|  |  |  |  |
| --- | --- | --- | --- |
| Mutation# | Strain (GenBank accession number) | Sample | Reference |
| C8A | NPA94b§ | NPA | this study |
| C10T | NPA94b | NPA | this study |
| T19C | Bld22d | blood | this study |
| T28C | NPA7aBld29bBld60c | NPAbloodblood | this studythis studythis study |
| T30A | NPA94b | NPA | this study |
| A33G | Bld29d | blood | this study |
| C36T | NPA44b | NPA | this study |
| T41C | Bld61d | blood | this study |
| T45C | Bld22c | blood | this study |
| C46T | NPA7a | NPA | this study |
| A57G | NPA7b | NPA | this study |
| T60C | NPA81d | NPA | this study |
| C74T | Bld22d | blood | this study |
| C77T | NPA1c | NPA | this study |
| T78C | NPA69c | NPA | this study |
| C85T | Bld38c | blood | this study |
| T93C | Bld61c | blood | this study |
| A99G | NPA89c | NPA | this study |
| A101T | NPA57a | NPA | this study |
| T111C | Bld61d | blood | this study |
| G116A | NPA81d | NPA | this study |
| T121A | Bld69c | blood | this study |
| G125A | HN057 (KC571691) | faeces | (Li et al., 2013) |
| A132-insAGGCGCTGCG-G133 | Brisbane 001 (EF520287)Brisbane 005 (FJ150410)CU-255 (EU358766) | NPANPANPA | (Bialasiewicz et al., 2007)(Bialasiewicz et al., 2009)(Payungporn et al., 2008) |
| T149A | Stockholm 380 (EF127908) | NPA | (Allander et al., 2007) |
| T149C | Stockholm 350 (EF127907)HN057 (KC571691) | NPAfaeces | (Allander et al., 2007)(Li et al., 2013) |
| T154C | NPA74cNPA81c | NPANPA | this studythis study |
| A162T | Bld37c | blood | this study |
| A163C | NPA89d | NPA | this study |
| A166G | NPA7d | NPA | this study |
| A173G | NPA81bBld61c | NPAblood | this studythis study |
| T180C | Brisbane 001 (EF520287)Brisbane 005 (FJ150410)Stockholm 380 (EF127908)Stockholm 350 (EF127907)HN057 (KC571691)CU-255 (EU358766)FZ52 (KM085447) | NPANPANPANPAfaecesNPANPA | (Bialasiewicz et al., 2007)(Bialasiewicz et al., 2009)(Allander et al., 2007)(Allander et al., 2007)(Li et al., 2013)(Payungporn et al., 2008)Xiu et al., unpublished |
| T180A | Brisbane 003 (EF520289) | NPA | (Bialasiewicz et al., 2007) |
| A181C | FZ52 (KM085447) | NPA | Xiu et al., unpublished  |
| T183A | Brisbane 001 (EF520287)Brisbane 005 (FJ150410)HN057 (KC571691)CU-255 (EU358766) | NPANPAfaecesNPA | (Bialasiewicz et al., 2007)(Bialasiewicz et al., 2009)(Li et al., 2013)(Payungporn et al., 2008) |
| T183G | Brisbane 003 (EF520289)FZ52 (KM085447) Stockholm 350 (EF127907)Stockholm 380 (EF127908) | NPANPANPANPA | (Bialasiewicz et al., 2007)Xiu et al., unpublished (Allander et al., 2007)(Allander et al., 2007) |
| G188A | NPA19aNPA1c | NPANPA | this studythis study |
| T189C | NPA94a | NPA | this study |
| A196G | Bld60d | blood | this study |
| C208T | Bld62c | blood | this study |
| T217A | NPA44a | NPA | this study |
| G224A | NPA19b | NPA | this study |
| A231G | NPA89a | NPA | this study |
| A234G | NPA69b | NPA | this study |
| T235A | NPA89c | NPA | this study |
| T247C | NPA19a | NPA | this study |
| T248A | NPA12g | NPA | this study |
| G290A | NPA89a | NPA | this study |
| G291A | Bld71c | blood | this study |
| T294C | Bld69c | blood | this study |
| T297C | NPA44dNPA81a | NPANPA | this studythis study |
| T300C | NPA44d | NPA | this study |
| A305G | NPA74f | NPA | this study |
| C306A | Brisbane 001 (EF520287)Brisbane 003(EF520289)Brisbane 005 (FJ150410)FZ52 (KM085447)Stockholm 380 (EF127908)Stockholm 350 (EF127907)HN057 (KC571691)CU-255 (EU358766) | NPANPANPANPANPANPAfaecesNPA | (Bialasiewicz et al., 2007)(Bialasiewicz et al., 2007)(Bialasiewicz et al., 2009)Xiu et al., unpublished (Allander et al., 2007)(Allander et al., 2007)(Li et al., 2013)(Payungporn et al., 2008) |
| T336C | NPA81a | NPA | this study |
| delG352 | NPA57b | NPA | this study |
| A371G | NPA12g | NPA | this study |
| A381T | Bld61c | blood | this study |
| T386C | Bld38d | blood | this study |
| A393G | Bld45d | blood | this study |
| T411C | Bld61d | blood | this study |
| A415G | Bld61c | blood | this study |
| T416C | NPA57b | NPA | this study |
| C418T | NPA89c | NPA | this study |
| C444T | NPA81b | NPA | this study |
| C455T | Bld61c | blood | this study |
| G484A | NPA1aNPA19aNPA69bNPA74dNPA74fNPA81a | NPANPANPANPANPANPA | this studythis studythis studythis studythis studythis study |

#Numbers refer to the nucleotides in the NCCR with the first nucleotide (nucleotide 1) adjacent to the start codon of the large T- and small t-antigens and the last nucleotide of the NCCR (nucleotide 513) just upstream of the startcodon of VP1.

§The letter refer to the clone analyzed after cloning the PCR product.